Asymptotic Behavior of Partition Functions with Graph Laplacian

O. Khorunzhiy

February 7, 2008

Abstract

We introduce the matrix sums that represent a discrete analog of the matrix integrals of random matrix theory. The summation runs over the set Γ_n of all possible n-vertex graphs γ_n weighted by $\exp\{-\beta \, \mathrm{Tr}\Delta_n\}, \beta > 0$, where $\Delta_n = \Delta(\gamma_n)$ is the analog of the Laplace operator determined on γ_n. Corresponding probability measure on Γ_n reproduces the well-known Erdős-Rényi ensemble of random graphs. Here it plays the same role as that played by the Gaussian Unitary Invariant Ensemble (GUE) in matrix models.

Regarding an analog of the matrix models with quartic potential, we study the cumulant expansion of related partition functions. We develop a diagram technique and describe the combinatorial structure of the coefficients of this expansion in two different asymptotic regimes $\beta = O(1)$ and $\beta = O(\log n)$ as $n \to \infty$.

Keywords: Laplace operator on graph, partition function, Erdős-Rényi random graphs, connected diagrams, Catalan numbers, Pólya equation.

1 Introduction

During last three decades, the studies of matrix models of theoretical physics have deeply influenced a number of branches of modern mathematics and mathematical physics. The central notion here is the partition function $Z_N(\beta, Q)$ given by the integral over the set $\mathcal{M}_N$ of all Hermitian N-dimensional matrices H

$$Z_N(\beta, Q) = \int_{\mathcal{M}_N} \exp\{-\beta \mathrm{Tr} H^2 + Q(H)\} dH = C_N \mathbf{E}^{(\beta)}_{\mathrm{GUE}}\{e^{Q(H)}\}, \qquad (1.1)$$

where Q is a "potential" function, dH is the Lebesgue measure, $C_N = Z_N(\beta, 0)$ is the normalizing constant, and $\mathbf{E}^{(\beta)}_{\mathrm{GUE}}\{\cdot\}$ denotes the mathematical expectation with respect to the probability measure $\mathcal{P}_N$ with the density

$$C_N^{-1} \exp\{-\beta \mathrm{Tr} H^2\} = C_N^{-1} \exp\{-\beta \sum_{i,j=1}^{N} |H_{ij}|^2\}. \qquad (1.2)$$

The measure $\mathcal{P}_N$ supported on $\mathcal{M}_N$ generates the Gaussian Unitary Invariant Ensemble of random matrices abbreviated by GUE (see monograph [12] for the detailed description of the ensemble and its properties).

The first non-trivial example of (1.1) is given by the quartic potential

$$Q(H) = \frac{g}{N}\mathrm{Tr}H^4. \tag{1.3}$$

The matrix model (1.1), (1.3) has served as the source of a series of deep results establishing connections between orthogonal polynomials, integrable systems, moduli spaces of curves and such combinatorial structures as maps (see [3] and [5] for the earlier and more recent results and references and [4] for the review). One of the simplest but important result is that the leading term of the formal asymptotic expansion of variable

$$\frac{1}{N^2}\log \mathbf{E}^{(\beta)}_{\text{GUE}}\left\{\exp\left(\frac{g}{N}\mathrm{Tr}H^4\right)\right\} \tag{1.4}$$

in the limit $N \to \infty$ is given by the series in powers of g with the coefficients determined by the numbers of 4-valent two-vertex maps dual to the famous quadrangulations of the compact Riemann manifold (see [19] for introductory description and references therein).

It should be noted that the GUE and its real symmetric and symplectic analogs represent very special class of random matrices. It is natural to ask about matrix models when the mathematical expectation in (1.4) is taken with respect to a measure different from that determined by GUE.

In present paper we introduce a discrete analog of (1.1), where the integration over $\mathcal{M}_N$ is replaced by the sum over the set Γ_n of all possible simple graphs γ_n with n vertices. In this setting the "kinetic energy" term $\mathrm{Tr}H^2$ is replaced by $\mathrm{Tr}\Delta_n$, where $\Delta_n = \Delta(\gamma_n)$ is the discrete analog of the Laplace operator determined on the graph γ_n. Corresponding Gibbs weight $\exp\{-\beta\ \mathrm{Tr}\Delta(\gamma_n)\}, \beta > 0$ generates the probability measure $\mu_n(\beta)$ on Γ_n.

A simple but non-trivial property of this ensemble plays a very important role in what follows. The observation is that the probability space (Γ_n, μ_n) coincides with the widely known Erdős-Rényi ensemble of random graphs with n vertices (see e.g. [2]). In this ensemble, the indicator functions of edges are represented by jointly independent Bernoulli random variables. As far as we know, this connection between Erdős-Rényi random graphs and the Gibbs measure $C_n \exp\{-\beta\ \mathrm{Tr}\Delta(\gamma_n)\}$ was not observed before.

Our aim is to consider the asymptotic behavior of $F_n = \log \mathbf{E}_{\mu_n}\{\exp(Q_n)\}$ where Q_n is given by the analog of quartic potential (1.3) and to explore the combinatorial structures that arise in this problem. We develop a diagram technique to study the cumulant expansion of F_n. We show that the leading terms of this expansion are related with the number of connected diagrams on the set of two-valent vertices. We derive recurrent relations for the numbers of such diagrams and describe the coefficients of the cumulant expansions of F_n in two different asymptotic regimes when $\beta = O(1)$ and $\beta = O(\log n)$ as $n \to \infty$. These recurrent relations generalise those for the Catalan numbers.

Corresponding generating function verifies an equation similar to the Pólya equation for the generating function of the rooted Cayley trees.

2 Graph Laplacian, matrix sums and random graphs

Given a finite graph with the set $V_n = \{v_1, \ldots, v_n\}$ of labelled vertices and $E_m = \{e^{(1)}, \ldots, e^{(m)}\}$ the set of simple non-oriented edges, the discrete analog of the Laplace operator $\Delta(\gamma)$ on graph γ can be determined as (see for example, [14]) by relation

$$\Delta(\gamma) = \partial^* \partial, \tag{2.1}$$

where ∂ is the difference operator determined on the space of complex functions on vertices $V_n \to \mathbf{C}$ and ∂^* is its conjugate determined on the space of complex functions on edges $E_m \to \mathbf{C}$.

It can be easily shown that in the canonical basis, the linear operator $\Delta(\gamma) = \Delta_n$ has $n \times n$ matrix with the elements

$$\Delta_{ij} = \begin{cases} \deg(v_j), & \text{if } i = j, \\ -1, & \text{if } i \neq j \text{ and } (v_i, v_j) \in E, \\ 0, & \text{otherwise,} \end{cases} \tag{2.2}$$

where $\deg(v)$ is the vertex degree. If one considers the $n \times n$ adjacency matrix $A = A(\gamma)$ of the graph γ,

$$A_{ij} = \begin{cases} 1, & \text{if } (v_i, v_j) \in E,\ i \neq j, \\ 0, & \text{otherwise,} \end{cases} \tag{2.3}$$

then one can rewrite the definition of Δ (2.2) in the form

$$\Delta_{ij} = B_{ij} - A_{ij}, \quad \text{where} \ \ B_{ij} = \delta_{ij} \sum_{l=1}^{n} A_{il}, \tag{2.4}$$

where δ is the Kronecker symbol

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$$

It follows from (2.1) that $\Delta(\gamma_n)$ has positive eigenvalues.

Let us consider the set Γ_n of all possible simple non-oriented graphs γ_n with the set $V = V_n$ of n labelled vertices. Obviously, $|\Gamma_n| = 2^{n(n-1)/2}$. Given an element $\gamma \in \Gamma_n$, it is natural to consider the trace $\mathrm{Tr}\Delta(\gamma)$ as the total energy of the graph γ. Then we can assign to each graph γ_n the Gibbs weight $\exp\{-\beta\ \mathrm{Tr}\Delta(\gamma_n)\}$, $\beta > 0$ and introduce the partition function

$$Z_n(\beta, Q) = \sum_{\gamma_n \in \Gamma_n} \exp\{-\beta\ \mathrm{Tr}\Delta_n + Q(\gamma_n)\}, \tag{2.5}$$

where $\Delta_n = \Delta(\gamma_n)$ and Q is an application $\Gamma_n \to \mathbf{R}$ that we specify later. Using the fact that $\mathrm{Tr}\Delta = \mathrm{Tr}\partial^*\partial$, one can consider (2.5) as a discrete analog of the partition function (1.1). We will see that this is especially interesting in the case of quartic potential (1.3).

Let us note that we should normalize the sum (2.5) by $|\Gamma_n|$, but this does not play any role with respect to our results. In what follows, we omit subscript n in Δ_n.

Definition (2.4) implies that

$$\mathrm{Tr}\Delta = \sum_{i=1}^{n} \Delta_{ii} = \sum_{i,j=1}^{n} A_{ij} = 2 \sum_{1\le i<j\le n} A_{ij}. \tag{2.6}$$

Then we can rewrite (2.5) in the form

$$Z_n(\beta, Q) = \sum_{\gamma_n\in\Gamma_n} e^{Q(\gamma_n)} \prod_{1\le i<j\le n} e^{-2\beta A_{ij}}. \tag{2.7}$$

It is easy to see that

$$Z_n(\beta, 0) = \left(1 + e^{-2\beta}\right)^{n(n-1)/2}. \tag{2.8}$$

Then the normalized partition function can be represented as

$$\hat{Z}_n(\beta, Q) = Z_n(\beta, Q)/Z_n(\beta, 0) = \mathbf{E}_\beta \left\{ e^{Q(\gamma)} \right\}, \tag{2.9}$$

where $\mathbf{E}_\beta\{\cdot\}$ denotes the mathematical expectation with respect to the measure supported on the set Γ_n. This measure assigns to each element $\gamma \in \Gamma_n$ the probability

$$P_n(\gamma) = \frac{e^{-2\beta|E(\gamma)|}}{(1 + e^{-2\beta})^{n(n-1)/2}},$$

where $E(\gamma)$ denotes the set of edges of the graph γ.

Given a couple (x, y), $x, y \in \{1, \dots, n\}$, one can determine a random variable a_{xy} on the probability space (Γ_n, P_n) that is the indicator function of the edge (v_x, v_y)

$$a_{xy}(\gamma) = \begin{cases} 1, & \text{if } (v_x, v_y) \in E(\gamma), \\ 0, & \text{otherwise.} \end{cases}$$

It is easy to show that the random variables $\{a_{xy}, 1 \le x < y \le n\}$ are jointly independent and are of the same Bernoulli distribution depending on β such that

$$a_{xy}^{(\beta)} = \begin{cases} 1, & \text{with probability } \dfrac{e^{-2\beta}}{1+e^{-2\beta}} = p, \\ 0, & \text{with probability } 1-p. \end{cases} \tag{2.10}$$

The probability space (Γ_n, P_n) is known as the Erdős-Rényi (or Bernoulli) ensemble of random graphs with the edge probability p. Since the series of pioneering papers by Erdős and Rényi, the asymptotic properties of graphs

(Γ_n, P_n), such as the size and the number of connected components, the maximal and minimal vertex degree and many others, are extensively studied (see [2, 7]). Spectral properties of corresponding random matrices A (2.3) and Δ (2.4) are considered in a series of papers (in particular, see [1, 6, 8, 9, 10, 13, 16]). In present paper we study the random graph ensemble (Γ_n, P_n) from another point of view motivated by the asymptotic behavior of partition functions (2.9) with certain "potentials" Q.

3 Partition functions and diagram technique

In the previous section, we have shown that the Gibbs weight $\exp\{-\beta \ \mathrm{Tr}\Delta(\gamma)\}$ generates the probability measure on graphs equivalent to that determined by the Erdős-Rényi ensemble of random graphs. This gives us an important tool for direct computation of averages of the form (2.9). Another important point is that this Gibbs weight leads us to the correct definition of discrete analogs of the matrix models (1.1) and in particular of the matrix models with quartic potentials (1.4). That is why one cannot neglect the Laplacian form of the Gibbs measure and start simply with computations of averages with respect to P_n.

3.1 Analog of the quartic potential

Let us determine the discrete analog of the partition function (1.1) with quartic potential (1.3). Once $\mathrm{Tr}H^2$ replaced by $\mathrm{Tr}(\partial^*\partial) = \mathrm{Tr}\Delta$, it is natural to consider

$$\mathrm{Tr}(\partial^*\partial\partial^*\partial) = \mathrm{Tr}\Delta^2$$

as the analog of $\mathrm{Tr}H^4$. Then the partition function (2.5) reads as

$$Z_n(\beta, g) = \sum_{\gamma_n \in \Gamma_n} \exp\{-\beta \ \mathrm{Tr}\Delta_n + g_n \mathrm{Tr}\Delta^2\}, \tag{3.1}$$

where g_n is to be specified. It follows from (2.3) and (2.4) that

$$\mathrm{Tr}\Delta^2 = \mathrm{Tr}B^2 + \mathrm{Tr}A^2 = \sum_{i,j=1}^{n} (A^2)_{ij} + \sum_{i,j=1}^{n} A_{ij}.$$

Then, using (2.6) and repeating computations of (2.7) and (2.8), we obtain representation

$$\hat{Z}_n(\beta, g) = Z_n(\beta, g)/Z_n(\beta, 0) = \left(\frac{1 + e^{-2\beta'}}{1 + e^{-2\beta}}\right)^{n(n-1)/2} \mathbf{E}_{\beta'}\{e^{g_n X_n}\}, \tag{3.2}$$

In this relation, we have denoted $\beta' = \beta - g_n$ and introduced the random variable

$$X_n = \sum_{i,j,l=1}^{n} a_{il} a_{lj}, \tag{3.3}$$

where a_{ij} are jointly independent random variables of the law (2.10) with β replaced by β'. The average $\mathbf{E}_{\beta'}$ denotes the corresponding mathematical expectation. In what follows, we omit the subscripts β and β' when they are not necessary. We study the limiting behavior of the cumulants of the random variable X_n in two different asymptotic regimes. The first limiting transition is determined by the choice $\beta_n = O(\log n)$ as $n \to \infty$. We study this case in Section 4. In Section 5, we consider the second asymptotic regime given by of $\beta_n = const$.

3.2 Cumulant expansion

Using the fact that a_{ij} are bounded by 1 and $X_n \leq n^3$, we can write that

$$\log \mathbf{E}\{e^{gX_n}\} = \sum_{k=1}^{\infty} \frac{g^k}{k!} Cum_k(X_n), \tag{3.4}$$

where $Cum_k(X_n)$ is the k-th cumulant of random variable X_n.

$$Cum_k(X_n) = \frac{d^k}{dg^k} \left(\log \mathbf{E}\{e^{gX_n}\} \right) |_{g=0}. \tag{3.5}$$

Denoting by Y_α the random variable $a_{il}a_{lj}$ with triplet $\alpha = (i,l,j)$, we can write that

$$Cum_k(X_n) = \sum_{\{\alpha_1,\ldots,\alpha_k\}} cum\{Y_{\alpha_1},\ldots,Y_{\alpha_k}\}, \tag{3.6}$$

where the sum runs over all possible values of $\alpha_s, s = 1,\ldots,k$ and

$$cum\{Y_{\alpha_1},\ldots,Y_{\alpha_k}\} = \frac{d^k}{dz_1 \cdots dz_k} \log \mathbf{E}\{\exp(z_1 Y_{\alpha_1} + \ldots + z_k Y_{\alpha_k})\}|_{z_r=0}. \tag{3.7}$$

The variable $cum\{Y_{\alpha_1},\ldots,Y_{\alpha_k}\}$ is also known in probability theory as the semi-invariant of k random variables Y.

Let us introduce a graphical representation of the set of variables Y. Given k values $\alpha_1,\ldots,\alpha_k$, we represent the set of random variables $Y_{\alpha_1},\ldots,Y_{\alpha_k}$ by the set of k labelled vertices $U = (u_1,\ldots,u_k)$. Regarding two variables Y_{α_r} and Y_{α_s}, we join corresponding vertices u_r and u_s by an edge $\varepsilon(r,s)$ if and only if Y_{α_r} and Y_{α_s} have at least one variable a in common. Considering all possible couples $(r,s), 1 \leq r < s \leq k$ and drawing corresponding edges, we obtain a graph that we denote by $G_k = (U_k, \mathcal{E}_k)$. It is clear that this graph $G_k = G_k(\vec{\alpha}_k)$ depends on particular value of the variable $\tilde{\alpha}_k = (\alpha_1,\ldots,\alpha_k)$. The following proposition is a well-known fact from the probability theory of random fields [11].

Lemma 3.1. *The semiinvariant $cum\{Y_{\alpha_1},\ldots,Y_{\alpha_k}\}$ is not equal to zero if and only if the graph $G_k(\tilde{\alpha}_k)$ is connected.*

Proof. Let us first note that if two random variables Y_{α_r} and Y_{α_s} have no variables a in common, then they are independent. Clearly, the same observation is true for the subfamilies of random variables Y. This means that if

the graph $G_k(\tilde{\alpha}_k)$ consists of two or more non-connected components, then the corresponding subsets of random variables Y are jointly independent. The characteristic property of the semiinvariants is that it vanishes in this case [11]. This completes the proof of Lemma 3.1.

Let us note that the fact that Y_{α_r} and Y_{α_s} have one or more variables a in common means that corresponding variables i, l, j coincide. In particular, the fact that Y_{α_r} and Y_{α_s} have exactly one variable a in common implies that one of the eight possibilities for the sets (i_r, l_r, j_r) and (i_s, l_s, j_s) occurs. For example, this happens when

$$(i_r = i_s, l_r = l_s, j_r \neq j_s) \quad \text{or} \quad (i_r = l_s, l_r = i_s, j_r \neq j_s). \tag{3.8}$$

We will refer to such cases as to the direct and inverse gluing, respectively.

Lemma 3.2. *Given β and g fixed, the number of terms in $Cum_k(X_n)$ is given by relation*

$$\#\{Cum_k(X_n)\} = O(n^{k+2}) \quad as\ n \to \infty. \tag{3.9}$$

Proof. To draw an edge $\varepsilon(r, s)$ of the graph G_k means to make equal at least two variables a taken from Y_{α_r} and Y_{α_s}, respectively. To make equal two or more a's means to make equal some of the variables (i_r, l_r, j_r) and (i_s, l_s, j_s).

Let us describe the process of drawing edges of G_k step by step. One starts with $3k$ variables i, j, l that can take values $1, \dots, n$ independently. Each gluing of two variables a diminishes by 2 the number of variables that move independently. To make the graph connected, we have to draw at least $k-1$ edges, so we are forced to perform at least $k-1$ gluings. When this is done, the number of variables that move independently is less or equal to $3k - 2(k-1) = k+2$. Maximizing the number of variables that can take different values, from 1 to n independently, we obtain

$$n(n-1)\cdots(n-(k-1)+1) = O(n^{k+2})$$

terms. Lemma 3.2 is proved.

3.3 Connected diagrams

Let us further develop the graphical representation of the set of variables Y and give more details for the description of corresponding connected graphs $\tilde{G}_k$. At this stage we make no difference between the direct and inverse gluings (3.8).

First let us note that each variable Y_α by itself can be represented by a graph of three vertices corresponding to variables i, l and j joined by two edges denoting random variables a_{il} and a_{lj}, respectively. Slightly modifying this, we can say that Y_{α_r} is represented by a vertex u_r with two off-spreads representing variables $a_{i_r l_r}$ to the left and $a_{l_r j_r}$ to the right from u_r. We say that the vertices u_r are two-valent. In $\tilde{G}_k$, there are k such two-valent vertices joined by $k-1$ edges that we refer now to as to arcs. The arcs join two different off-spreads

that we glue during the procedure described in the proof of the Lemma 3.2. The family of k two-valent vertices together with $k-1$ arcs represent *a diagram* that we denote by δ_k. This diagram provides more information than the graph $\tilde{G}_k$ because it shows exactly which off-spreads are glued between themselves. If one forgets the two-valent structure of u_r's, one gets the tree $\tilde{G}_k$. One of the possible example of δ_k and corresponding $\tilde{G}_k$ for $k=5$ is given on figure 1.

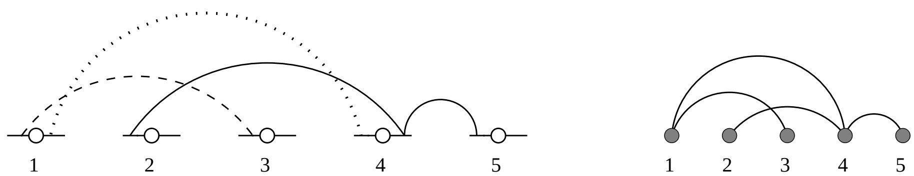

Figure 1: Connected diagram δ_k and corresponding tree $\tilde{G}_k$ for $k=5$.

One can see that several off-spreads can be glued together. In this case we say that they are colored by the same color. We call the off-spreads that remain non-glued as the free off-spreads and leave them non-colored (or grey). On figure 1 the diagram δ_5 contains three grey elements and three color groups of 2, 2, and 3 elements.

Clearly, one can draw several diagrams that represent the same coloring of the off-spreads. An example of such two diagrams δ_k and δ'_k with $k=5$ is given on figure 2. We call the diagram *the reduced diagram* when the off-spreads of the same color are joined by the arcs connecting the nearest neighbors. On figure 2 the diagram δ_k is the reduced one, the diagram δ'_k is not.

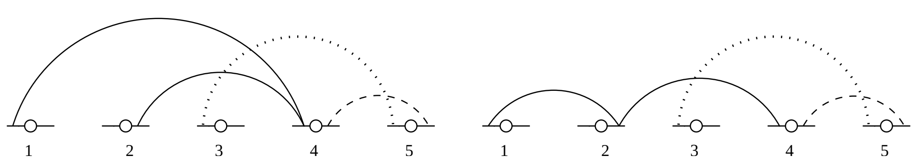

Figure 2: Non-reduced diagram δ'_k and the equivalent reduced one.

This argument explains the difference between the diagrams we have and the set of trees on labelled vertices. Indeed, the diagrams δ_k and δ'_k are equivalent but the corresponding trees are not.

With this diagram representation, we see that the leading contribution to (3.9) comes from the family $\mathcal{D}_k$ of connected acyclic reduced diagrams δ_k drawn on the set of k two-valent vertices with ordered off-spreads. The number of such diagrams $d_k = |\mathcal{D}_k|$ is determined in the next section. We complete this section with the following simple proposition.

Lemma 3.3. *Let us consider a connected reduced diagram δ_k that have r color groups of arcs, with $\mu_1, \mu_2, \ldots, \mu_r$ arcs in each color group. Then*

$$\mu_1 + \mu_2 + \ldots + \mu_r = k-1. \tag{3.10}$$

There are $k-r+1$ grey elements in δ_k.

Proof. Each diagram generates a tree on k vertices $u_1,\dots,u_k$. So the total number of arcs is equal to $k-1$. The group of μ_s arcs of the same color produces μ_s+1 color elements. The total number of colored elements is $k-1+r$. Then the number of non-colored (grey) elements in δ_k is $2k-(k-1+r)=k-r+1$.

4 Sparse random graphs

In present section we study the case when the "temperature " $T_n = 1/\beta_n$ vanishes when $n\to\infty$

$$\beta_n = \frac{1}{2}\log\left(\frac{n}{c}\right)(1+o(1)), \quad n\to\infty \tag{4.1a}$$

with some $c>0$. This corresponds to the random graph ensemble (3.2) with vanishing edge probability

$$p_n = \left(\frac{\bar c}{n}\right)^{(1+o(1))}, \quad n\to\infty, \tag{4.1b}$$

where we denoted $\bar c = ce^{2g(n,c)}$ and $g(n,c)=g/c$. In this paper we consider the limiting transition when $n,c\to\infty$ and $c=o(n)$. The case of $n\to\infty, c=const$ will be studied in separate publication. The main results of this section are presented by the following two statements.

Theorem 4.1.
Given $k\ge 2$, there exists the limit

$$\lim_{n,\bar c\to\infty,\bar c=o(n)} \frac{1}{n\bar c^{k+1}} Cum_k(X_n) = 2^{k-1}d_k, \tag{4.2}$$

where the numbers $d_k, k\ge 2$ are determined by following recurrent relation

$$d_k = 2kd_{k-1} + \sum_{j=1}^{k-2}\binom{k-1}{j}(j+1)(k-j)\ d_j\,d_{k-1-j}\ ,\ k\ge 3 \tag{4.3}$$

with the initial condition $d_2=4$.

Let us consider the auxiliary numbers

$$h_k = \frac{(k+1)d_k}{k!} \quad \text{for} \quad k\ge 2.$$

It is easy to deduce from (4.3) that the sequence h can be determined by the following recurrent relation

$$h_k = \frac{k+1}{k}\sum_{j=0}^{k-1} h_j\ h_{k-1-j}, \quad h_0 = 1. \tag{4.4}$$

Proposition 4.2.
The generating function $h(z) = \sum_{k=0}^{\infty} h_k z^k$ *is determined in the complex domain* $R_{1/8} = \{z \in \mathbf{C} : |z| < 1/8\}$ *and verifies there the following equation*

$$h(z) = \exp\{2zh(z)\}. \tag{4.5}$$

It follows from (4.5) that

$$h_k = 2^k \frac{(k+1)^{k-1}}{k!}, \quad k \geq 1. \tag{4.6}$$

Using (4.4), it is easy to find the first values of d

$$d_2 = 4, \; d_3 = 32, \; d_4 = 400, \; d_5 = 6912, \; d_6 = 153664.$$

The general form of $\{d_k\}$ follows from (4.6)

$$d_k = 2^k (k+1)^{k-2}, \quad k \geq 2. \tag{4.7}$$

Equation (4.5) is similar to the Pólya equation for the generating function of the rooted trees on labelled vertices [15, 17] (see also relation (4.17) below). The explicit expression (4.6) resembles the number of Cayley trees on $k+1$ vertices. However, we did not find any obvious one-to-one correspondence between the set of Cayley trees on $k+1$ vertices and the k-vertex diagrams we count. While the skeletons of our diagrams are given by trees, the maximal degree of these trees is bounded by 4. Also, there are equivalent diagrams that have different tree skeletons. This makes the set of the diagrams we study quite different from the family of Cayley trees. Relation (4.4) generalizes recurrent relations for the Catalan numbers. Up to our knowledge, the class of diagrams $\mathcal{D}_k$ as well as the numbers d_k were not considered previously.

Proof of Theorem 4.1. Let us separate the set of all possible values of triplets $\alpha_1, \ldots, \alpha_k$ containing n^{3k} elements into the classes of equivalence labelled by $\delta_k \in \mathcal{D}_k$. This is done in obvious way. Then we can write that

$$\sum_{\alpha_1,\ldots,\alpha_k} cum\{Y_{\alpha_1}, \ldots, Y_{\alpha_k}\} = \sum_{\delta_k \in \mathcal{D}_k} \mathcal{N}(\delta_k) \; cum\{\bar{Y}(\delta_k)\}(1 + o(1)), \quad n \to \infty \tag{4.8}$$

where $\mathcal{N}(\delta_k)$ denotes the number of elements in the equivalence class labelled by δ_k,

$$cum\{\bar{Y}(\delta_k)\} = cum\{Y_{\bar{a}(\delta_k)_1}, \ldots, Y_{\bar{a}(\delta_k)_k}\},$$

and $\bar{a}(\delta_k)$ is one of the representative of this equivalence class. For instant, we can choose $\bar{a}(\delta_k) = (\alpha_1, \ldots, \alpha_k)$ with minimal possible values of i_1, l_1, j_1, i_2, and so on, in the way such that $(\alpha_1, \ldots, \alpha_k)$ belongs to the class δ_k. It follows from Lemma 3.2 that

$$\mathcal{N}(\delta_k) = n(n-1)\cdots(n-k-1) = n^{k+2}(1 + o(1)), \quad n \to \infty. \tag{4.9}$$

Using the basic property of the semi-invariants, we can write that [11]

$$cum\{\bar{Y}(\delta_k)\} = \sum_{\pi_k} \mathbf{E}\{\tilde{Y}_{T_1(\pi_k;\bar{\alpha}(\delta_k))}\} \cdots \mathbf{E}\{\tilde{Y}_{T_\sigma(\pi_k;\bar{\alpha}(\delta_k))}\}(-1)^{\sigma-1}(\sigma-1)!\,, \quad (4.10)$$

where the sum runs over all unordered partitions $T_1, \dots, T_\sigma$ of the set of vertices $U_k = \{u_1, \dots, u_k\}$; that is over all families of non-empty non-intersecting subsets T_s of U_k giving in sum all U_k. The second argument of T's reminds us that the variables $\bar{\alpha}(\delta_k)_s$ attached to u_s are chosen according to the rule prescribed by δ_k. Random variable $\tilde{Y}_{T_s(\pi_k;\bar{\alpha}(\delta_k))}$ is given by the product of corresponding random variables Y.

Let us consider the term of the sum (4.10) that corresponds to the trivial partition $\pi_k^{(1)}$ of U_k consisting of one subset only: $T_1 = U_k$. This term is given by the average value

$$\mathbf{E}\{Y_{\bar{a}(\delta_k)_1} \cdots Y_{\bar{a}(\delta_k)_k}\} = W(\pi_k^{(1)}).$$

According to Lemma 3.3, there are $k - r + 1$ non-colored elements in diagram δ_k, where r is the number of color groups of arcs in δ_k. These grey elements represent independent random variables a that are also independent from elements belonging to color groups. Then we obtain the factor

$$(\mathbf{E}a)^{k-r+1} = \left(\frac{\bar{c}}{n}\right)^{k-r+1}.$$

Each color group produces the factor $\mathbf{E}a^{\mu_s+1} = \mathbf{E}a = \bar{c}/n$ and there are r color groups. Then the weight of the partition $\pi_k^{(1)}$ is

$$W(\pi_k^{(1)}) = \left(\frac{\bar{c}}{n}\right)^{k+1}. \quad (4.11)$$

Now it is clear that any other partition π_k produces the term of the order $o((\bar{c}/n)^{k+1})$ that is evidently smaller than that of (4.10). More precisely, we can write that

$$\mathbf{E}\{\tilde{Y}_{T_1(\pi_k;\bar{\alpha}(\delta_k))}\} \cdots \mathbf{E}\{\tilde{Y}_{T_\sigma(\pi_k;\bar{\alpha}(\delta_k))}\} = \left(\frac{\bar{c}}{n}\right)^{k+1+\chi}, \quad (4.12)$$

where χ is equal to the number of the arcs θ of δ_k such that the left and right feet of θ belong to different subsets T' and T'' of the partition π_k under consideration. We say that these arcs are cut by partition π_k. In particular, the partition $\pi_k^{(k)}$ of U_k into k subsets produces the factor

$$W(\pi_k^{(k)}) = \left(\frac{\bar{c}}{n}\right)^{2k}$$

because all $k-1$ arcs of δ_k are cut by this partition.

Remembering that each arc can be drawn in the direct and inverse sense (3.8) and gathering formulas (3.6), (4.8), (4.9) and (4.11), we conclude that

$$\frac{1}{n\bar{c}^{k+1}} Cum_k(X_n) = 2^{k-1} d_k (1 + o(1)), \quad 1 \ll \bar{c} \ll n,$$

where $d_k = |\mathcal{D}_k|$ is the total number of diagrams δ_k. Let us derive recurrent relations for d_k.

Let us consider one particular diagram δ_k and denote by $P_j(\delta_k)$ some part of δ_k that consists of certain j vertices and all arcs that join them.

Let us now consider the last vertex u_k. The following two cases are possible:

(a) there is only one arc that joins u_k with $P_{k-1}(\delta_k)$ and

(b) there are two arcs that join u_k with $P_{k-1}(\delta_k)$.

Clearly, the case when u_k is joined with $P_{k-1}(\delta_k)$ by three or more arcs is prohibited because of the absence of cycles in corresponding tree $\tilde{G}_k$.

In the first case $P_{k-1}(\delta_k)$ is a connex diagram. One can join this diagram with u_k by an arc that have the left foot supported either on grey element or on the maximal element of the color group. If there are r colour groups in $P_{k-1}(\delta_k)$, then there are $k-r$ grey elements (see Lemma 3.3). There are r maximal color elements. Thus one can choose k elements to put the left foot of the arc. The right foot can be put on one of the two elements of u_k. Clearly, $|P_{k-1}| = d_{k-1}$. So, the case (a) produces $2kd_{k-1}$ connected diagrams. On figure 3 we illustrate this situation.

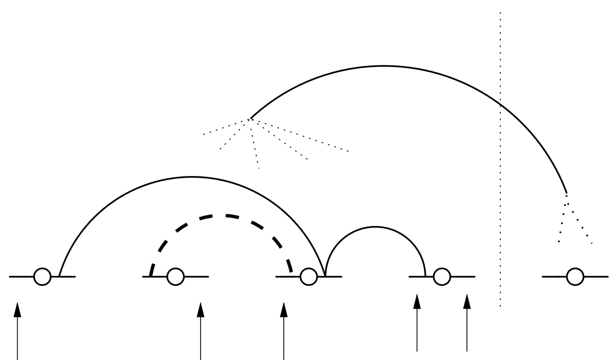

Figure 3: Arc to join u_k with P_{k-1}. Arrows show possible emplacements of the left foot.

Not let us pass to the case (b). In this case $P_{k-1}(\delta_k)$ is splitted in two connected diagrams. Let us assume that the left element of u_k is connected by an arc with the diagram constructed on j vertices. There are $j+1$ possibilities to put the left foot of the arc. The right element of u_k joined to the component of $k-1-j$ elements. There are $k-j$ possibilities to do this. The choice of the vertices to produce the component of j elements gives $\binom{k-1}{j}$ possibilities. Then we get the formula (4.3). It is easy to see that on the way described we obtain all the diagarms of $\mathcal{D}_k$.

Theorem 4.1 is proved.

Proof of Proposition 4.2. Recurrent relation (4.4) resembles very much the recurrent relation determining the moments of the famous semi-circle distribu-

tion from random matrix theory

$$m_k = v^2 \sum_{j=0}^{k-1} m_j m_{k-1-j}, \quad m_0 = 1 \tag{4.13}$$

with a parameter $v > 0$ [18]. It is also known that the numbers $m_k = m_k(v)$ are proportional to the Catalan numbers C_k;

$$m_k(v) = v^{2k} C_k = \frac{1}{k+1}\binom{2k}{k}. \tag{4.14}$$

One can easily deduce from (4.13) that

$$m_k \le (2v)^{2k}.$$

Comparing (4.4) with (4.13) taken for $v^2 = 2$, we conclude that

$$h_k \le 8^k \tag{4.15}$$

for all $k \ge 1$. This proves regularity of the generating function $h(z)$ in the domain $z \in R_{1/8}$.

Now let us show that $h(z)$ verifies equality (4.5). It is easy to see that $h(x), -1/8 < x < 1/8$ is determined by the differential equation

$$h'(x) = \frac{2h^2(x)}{1 - 2xh(x)}, \quad h(0) = 1. \tag{4.15}$$

Indeed, rewriting (4.4) in the form

$$h_k = \sum_{j=0}^{k-1} h_j \ h_{k-1-j} + \frac{1}{k} \sum_{j=0}^{k-1} h_j \ h_{k-1-j}, \tag{4.16}$$

we multiply both parts of (4.16) by x^k and after summation over $k \ge 1$, we get relation

$$h(x) - 1 = xh^2(x) + \int_0^x h^2(y) dy.$$

Passing to the derivatives, we get equation (4.15).

With the help of the substitution

$$xh(x) = \psi(x)$$

we reduce (4.15) to equation

$$\psi' = \frac{\psi}{x(1 - 2\psi)}$$

that gives $\psi e^{-2\psi} = Cx$. Observing that $C = 1$, we get (4.5).

To determine the explicit form of coefficients h_k, we use the standard technique of the contour integration [17]. First let us note that function $\psi(z) = zh(z)$ verifies the Pólya equation

$$\psi(z) = ze^{2\psi(z)}$$

and that the inverse function $\psi^*(w) = we^{-2w}$ is regular in the vicinity of the origin. By the Cauchy formula, we have

$$\psi_k = \frac{1}{2\pi \ \mathrm{i}} \oint \frac{\psi(z)}{z^{k+1}} dz.$$

Changing variables by $z = \psi^*(w)$, we get $dz = (1 - 2w)e^{-2w}dw$ and find that

$$\psi_k = \frac{1}{2\pi \ \mathrm{i}} \oint \frac{1 - 2w}{w^k} e^{2wk} dw.$$

Then

$$\psi_k = 2^{k-1} \frac{k^{k-2}}{(k-1)!}$$

and (4.6) follows. Proposition 4.2 is proved.

Accepting that $d_1 = 1$ and introducing the exponential generating function

$$D(\tau) = \sum_{k=1}^{\infty} \frac{2^{k-1} d_k}{k!} \tau^k, \tag{4.17}$$

and taking into account (4.2), one can write formally for the partition function (3.2) that

$$\lim_{n\to\infty} \frac{1}{nc} \log \hat{Z}_n(\beta, \frac{g}{c}) \simeq \frac{e^{2g}}{2} D(ge^{2g}). \tag{4.18}$$

The rigorous derivation of this equality will be done in separate publication.

5 The case of constant edge probability

Now it is easy to prove the following statement that concerns partition function (3.2), where X_n is determined by (3.3).

Proposition 5.1. *Given $k \geq 1$, there exists the limit*

$$\lim_{n\to\infty} \frac{1}{n^2} Cum_k(\frac{g}{n} X_n) = g^k \sum_{\pi_k} W(\pi_k; p) = g^k \sum_{\delta_k \in \mathcal{D}_k} w(\delta_k; p), \tag{5.1}$$

where W and w are polynomials in p involving degrees $p^{k+1}, \dots, p^{2k}$ and $p = e^{-2\beta'}/(1 - e^{-2\beta'})$ with $\beta' = \beta - g$.

Proof. We follow the lines of the proof of Theorem 4.1 because the formulas (4.8), (4.9) and (4.10) are still valid in the present case with obvious changes.

Regarding (4.10), let us consider the term that corresponds to the trivial partition $\pi_k^{(1)}$. It is easy to see that the formula (4.11) reads as

$$W(\pi_k^{(1)}) = (\mathbf{E}a)^k \mathbf{E}a^k = p^{k+1}.$$

Clearly, relation (4.12) takes the form

$$\mathbf{E}\{\tilde{Y}_{T_1(\pi_k;\bar{\alpha}(\delta_k))}\} \cdots \mathbf{E}\{\tilde{Y}_{T_\sigma(\pi_k;\bar{\alpha}(\delta_k))}\} = p^{k+1+\chi},$$

and this weight does not vanish as it was before. In particular, another trivial partition of the set U_k into k subsets produces the factor p^{2k}. Relation (4.9) completes the proof of the first equality of (5.1).

The second equality in (5.1) represents the another order of the summation: we fix a partition π_k and consider the sum over all diagrams $\delta_k \in \mathcal{D}_k$

$$w(\delta_k;p) = (-1)^{\sigma-1}(\sigma-1)! \sum_{\delta_k \in \mathcal{D}_k} \mathbf{E}\{\tilde{Y}_{T_1(\pi_k;\bar{\alpha}(\delta_k))}\} \cdots \mathbf{E}\{\tilde{Y}_{T_\sigma(\pi_k;\bar{\alpha}(\delta_k))}\}. \quad (5.2).$$

Certainly, the numbers $W(\pi_k;p)$ and $w(\delta_k;p)$ are uniquely determined. It is possible to obtain recurrent relation that determine the weight $W(\pi_k;p)$. This recurrent relation generalize (4.3) but it is very cumbersome and complicated. We do not present it here.

6 Discussion and perspectives

We have introduced the discrete analog of the matrix models with quartic potentials. We have shown that in this approach the Erdős-Rényi ensemble naturally arises and plays the same role as that played by GUE for the matrix models. Regarding cumulant expansion of the normalized partition function, we have shown that the connected diagrams on two-valent vertices replace the four-valent two-vertex maps seen in analogous situation in matrix models. Let us discuss our results with respect to the properties of random graphs.

Given a graph γ with adjacency matrix $A(\gamma)$, the variable

$$\frac{1}{2}X_n = \frac{1}{2}\sum_{i,j=1}^{n} (A^2)_{ij} = \frac{1}{2}\sum_{i,l,j=1}^{n} A_{il}A_{lj} \quad (6.1)$$

represents the number of all possible two-step walks over γ.

From this point of view, it is natural to ask the same question about asymptotic behavior of the q-step walks and study the terms of the cumulant expansion of variable

$$\Theta_q^{(n)}(c;g) = \frac{1}{n}\log \mathbf{E}_{\beta'}\left\{e^{gX_n^{(q)}}\right\}, \quad \text{with} \quad X_n^{(q)} = \sum_{i,j=1}^{n} (A^q)_{ij}, \quad q \geq 2. \quad (6.2)$$

It is not hard to show that in the situation of (6.2), the reasonings of Sections 3 and 4 remains true with obvious changes. The first modification is that instead

of the two-valent vertices u_s we have to consider q-valent vertices. More precisely, u_s are represented by the linear graph with $q+1$ subvertices and q edges that join them. But it is not hard to see that the number of corresponding diagrams is equivalent to the number of diagrams constructed on the set of q-valent vertices, or in other words q-stars with ordered (or labelled) off-spreads. Let us count the number $d_k^{(q)}$ of connected reduced acyclic diagrams on q-stars.

We start with the case $q = 3$. It is not difficult to repeat the proof of Proposition 4.2 and to derive recurrent relations for $k \geq 2$

$$d_k^{(3)} = 3(2k-1)\, d_{k-1}^{(3)} + 3I_{\{k\geq 3\}} \times \left(\sum_{j_1+j_2=k-1, j_i\geq 1} \frac{(k-1)!}{j_1!j_2!} (2j_1+1)(2j_2+1) d_{j_1}^{(3)} g_{j_2}^{(3)} \right)$$

$$+ I_{\{k\geq 4\}} \times \left(\sum_{j_1+j_2+j_3=k-1, j_i\geq 1} \frac{(k-1)!}{j_1!j_2!j_3!} (2j_1+1)(2j_2+1)(2j_3+1) d_{j_1}^{(3)} d_{j_2}^{(3)} d_{j_3}^{(3)} \right) \tag{6.3}$$

with the initial condition $d_1^{(3)} = 1$. Introducing the auxiliary numbers

$$h_k^{(3)} = (2k+1) d_k^{(3)} / k!$$

and setting $h_0^{(3)} = 1$, it is not hard to show that (6.3) is equivalent to the following recurrent relation

$$h_k^{(3)} = \frac{2k+1}{k} \sum_{j_1+j_2+j_3=k-1, j_i\geq 0} h_{j_1}^{(3)} h_{j_2}^{(3)} h_{j_3}^{(3)}, \quad k \geq 1. \tag{6.4}$$

Using (6.4), we derive differential equation

$$\frac{dh^{(3)}(x)}{dx} = \frac{3\left(h^{(3)}(x)\right)^3}{1 - 6x\left(h^{(3)}(x)\right)^2}, \quad h^{(3)}(0) = 1,$$

where $h^{(3)}(x) = \sum_{k\geq 0} h_k^{(3)} x^k$.

In the general case of $q \geq 2$, we obtain equation

$$\frac{dh^{(q)}(x)}{dx} = \frac{q\left(h^{(q)}(x)\right)^q}{1 - (q^2-q)x\left(h^{(q)}(x)\right)^{q-1}}, \quad h^{(q)}(0) = 1. \tag{6.5}$$

Substitution $\psi(x) = x\left(h^{(q)}(x)\right)^{q-1}$ leads us to equation

$$\psi'(x) = \frac{1}{x} \cdot \frac{\psi(x)}{1 - (q^2-q)\psi(x)}.$$

Resolving it and returning back to the function $h(x) = h^{(q)}(x)$, we arrive at the equation that generalizes the Pólya equation

$$h(x) = \exp\{qxh^{q-1}(x)\}, \quad q \geq 2. \tag{6.6}$$

Using again the function ψ and repeating computations of Section 4, we obtain relations to determine the coefficients $h_k^{(q)}, k \ge 0$;

$$\sum_{j_1+\ldots+j_{q-1}=k} h_{j_1}\cdots h_{j_{q-1}} = \psi_{k+1} = (q^2-q)^k \frac{(k+1)^{k-1}}{k!}, \quad h_0 = 1.$$

In particular, for $q = 3$, we obtain the first values $h_1^{(3)} = 3$, $h_2^{(3)} = 45/2$, $h_3^{(3)} = 1071/6$ that correspond to the sequence $d_1^{(3)} = 1$, $d_2^{(3)} = 9$, $d_3^{(3)} = 153$.

As we have seen, the discrete matrix model (2.5) we proposed represents an interesting source of questions about corresponding combinatorial structures. The limiting expressions we compute represent the first, zero-order approximation to the free energy per site of this model. It could be interesting to further develop the diagram approach to study the next terms of the $1/n$-expansion of this free energy and justify equality (4.18). A special attention is to be paid for the limiting transition $n \to \infty$, $c = const$. One can expect that the combinatorial structure of the cumulants of X_n is determined again by the diagrams of the type $\mathcal{D}_j$ but now all of the diagrams with $j = 1, \ldots, k$ are to be involved. This is a subject of a separate publication.

References

[1] M. Bauer, O. Golinelli, Random incidence matrices: moments of the spectral density, *J. Stat. Phys.* **103** (2001) 301-337

[2] B. Bollobás. *Random Graphs*, Cambridge University Press, Cambridge, 2001.

[3] D. Bessis, C. Itzykson, J.-B. Zuber, Quantum field theory techniques in graphical enumeration, *Adv. Appl. Math.* **1** (1980) 109-157

[4] P. Di Francesco. 2D quantum gravity, matrix models and graph combinatorics. *Applications of random matrices in physics*, NATO Sci. Ser. II Math. Phys. Chem., **221**, Springer, Dordrecht, 2006 pp. 33-88

[5] N. M. Ercolani and K. D. T.-R McLaughlin, Asymptotics of the partition function for random matrices via Riemann-Hilbert techniques and applications to graphical enumeration. *Int. Math. Res. Not.* (2003) no. 14, 755–820

[6] S. Evangelou, A numerical study of sparse random matrices, *J. Stat. Phys.* **69** (1992) 361-383

[7] S. Janson, T. Łuczak, and A. Rucinski. *Random graphs.* Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[8] O. Khorunzhiy, W. Kirsch, and P. Müller, Lifshits tails for spectra of Erdős-Rényi random graphs. *Ann. Appl. Probab.* **16** 295-309 (2006)

[9] A. Khorunzhy, M. Shcherbina, and V. Vengerovsky, Eigenvalue distribution of large weighted random graphs. *J. Math. Phys.* **45** (2004) 1648-1672

[10] M. Krivelevich and B. Sudakov, The largest eigenvalue of sparse random graphs, *Combin. Probab. Comput.* **12** (2003) 61-72

[11] V. Malyshev and R. Minlos. *Gibbs random fields. Cluster expansions.* Mathematics and its Applications (Soviet Series), **44**. Kluwer Academic Publishers Group, Dordrecht, 1991.

[12] M.L. Mehta. *Random Matrices.* Academic press, 1991

[13] A.D. Mirlin and Ya. V. Fyodorov, Universality of level correlation function of a sparse random matrix, J. Phys. A **24** (1991) 2273-2286

[14] B. Mohar. *The laplacian spectrum of graphs, graph theory, combinatorics, and applications.* Wiley, New York, 1991

[15] R. Ottes, The number of trees, *Ann. Math.* **49** (1948) 583-599

[16] G.J. Rodgers and A.J. Bray, Density of states of sparse random matrix, *Phys. Rev. B* **37** (1988) 3557-3562

[17] K. Rybakov, A. Menshikov et al. *Combinatorial Analysis: Problems and Exercises*, Moscow, Nauka (1982) (in russian)

[18] E. P. Wigner, Characteristic vectors of bordered matrices with infinite dimensions. *Ann. of Math.* **62** (1955) 548–564.

[19] A. Zvonkin, Matrix integrals and map enumeration: an accessible introduction. *Math. Comput. Modelling* **26** (1997) 281–304